

WO 01/47274 A1

Fine granular scalable video with embedded DCT coding of the enhancement layer

Background of the Invention

The present invention generally relates to video compression, and more particularly to a method for encoding enhancement layer video data in an embedded fashion in order to achieve fine granular scalable video.

5 Scalable video coding is a desirable feature for many multimedia applications and services that are used in systems employing decoders with a wide range of processing power. Several types of video scalability schemes have been proposed such as temporal, spatial and quality scalability. All of these types consist of a base layer and an enhancement layer. The base layer is the minimum amount of data required to decode the video stream,
10 while the enhancement layer is the additional data required to provide an enhanced video signal.

For each type of video scalability scheme, a particular scalability structure is defined. One type of structure is known as fine granularity scalability (FGS), which has been proposed and will soon become part of the MPEG-4 multimedia standard. The use of FGS
15 primarily targets applications where video is streamed over heterogeneous networks in real time. Further, FGS enables the bandwidth to be adapted by encoding content once for a range of different bit rates, which enables a video transmission server to change the transmission rate dynamically without in depth knowledge of or parsing the video stream.

Currently, there is an implementation of the proposed FGS structure in
20 MPEG-4 as a reference for the core experiment on this standardization activity. This particular implementation uses the current MPEG-4 coding standard as the base layer encoding scheme. The MPEG-4 implementation also encodes the enhancement layer as the difference between the discrete cosine transform (DCT) coefficients of the original picture and the reconstructed base layer DCT coefficients. Further, the enhancement coding scans
25 through the difference (or residual) DCT coefficients bit-plane by bit-plane to encode a series of 1's and 0's as a refinement of the base layer DCT coefficients.

One major limitation of the above-described implementation is that the enhancement layer encoder scans each individual bit plane of the residual DCT coefficients from the most significant to the least significant bit, block by block. In other words, for each

bit plane, a whole DCT coefficient block is scanned before subsequent blocks are scanned. Thus, this requires coding of one bit-plane of all of the DCT coefficients for the whole image in order to refine the entire picture. Therefore, the enhancement layer bit stream generated by this implementation contains only a limited number of scalability layers.

5 Embedded or progressive coding of still images was first utilized for wavelet image coding, which was later extended to DCT image coding. Thus, embedded DCT coding algorithms have been proposed in the past. These coding algorithms retained high compression efficiency while achieving high scalability in the resulting bit streams. Therefore, these algorithms may be alternatives for the FGS encoding structure.

10

Summary of the Invention

The present invention is directed to a method for encoding video data in an embedded fashion in order to achieve fine granular scalable video. The method according to the present invention still scans the DCT coefficients bit-plane by bit-plane. However, the present invention differs in that it incorporates DCT frequency domain scanning besides spatial and bit-plane scanning.

15

The method according to the present invention includes the video data being transformed into a plurality of DCT coefficients. Further, the DCT coefficients are arranged into sub-groups and the DCT coefficients are scanned according to the sub-groups. The DCT coefficients being scanned by the sub-groups enables a higher level of scalability to be achieved.

20

Brief Description of the Drawings

Referring now to the drawings were like reference numbers represent corresponding parts throughout:

25

Figure 1 is a diagram showing one example of the sub-grouping of the DCT coefficients according to the present invention;

Figure 2 is a diagram of the pseudo code for the DCT coefficient scanning according to the present invention; and

30

Figure 3 is a diagram showing one example of the DCT coefficient scanning according to the present invention.

Detailed Description

The present invention is directed to a method of encoding enhancement layer video data in order to achieve fine granular scalable video. This is accomplished by scanning enhancement layer DCT coefficients in an embedded fashion. The method according to the present invention still scans the DCT coefficients bit-plane by bit-plane. However, the present invention differs in that it incorporates DCT frequency domain scanning besides spatial and bit-plane scanning.

As previously described, the enhancement layer encoder proposed for MPEG-4 scans the DCT coefficients block by block for each bit-plane. This means that every DCT coefficient in a block is scanned before proceeding to the next block of coefficients. In contrast, the present invention scans a subset of the DCT coefficients included in each of the blocks before proceeding to the next block for each bit-plane. In other words, for each bit-plane, the DCT coefficient blocks are scanned a portion at a time. Thus, when the present invention finishes scanning one sub-group for all of the DCT coefficient blocks included in a bit-plane, it will then scan the next sub-group of DCT coefficients in the same bit plane. The DCT coefficients being scanned by the sub-groups enables a higher level of scalability to be achieved.

It is preferred that method of scanning the DCT coefficients according to the present invention is performed in two steps. These two steps include arranging the DCT coefficients of each bit plane into sub-groups and then scanning each bit-plane according to these sub-groups. However, prior to this scanning, other processing is also performed that is also common to the enhancement layer encoder proposed for MPEG-4. This processing includes generating the enhancement layer DCT coefficients from the video data. Further, the enhancement layer DCT coefficients are digitized and then divided into bit-planes.

In the first step of the method according to the present invention, the DCT coefficients of each bit plane are arranged into sub-groups. Since each bit-plane includes a number of DCT coefficient blocks, the first step produces DCT coefficient blocks having a similar sub-group arrangement. In one example, the DCT coefficients are arranged into four sub-groups having a pyramid-like structure that resembles the wavelet transform, as shown in Figure 1.

As can be seen from Figure 1, each of the sub-groups 66,68,70,72 is shaped in an L-configuration and is characterized by a distinctive gray level. Further, it is also evident that the first sub-group 66 includes just the DC coefficient, the second sub-group 68 includes the three neighboring DCT coefficients, the third sub-group 70 includes the next twelve coefficients, etc.

It should be noted that the arrangement of the DCT coefficients shown in Figure 1 represents only one example. According to the present invention, other arrangements are contemplated. For example, the first and second sub-groups 66,68 can be combined into a one sub-group. Further, the configuration of the sub-groups can be changed.

5 Instead of the L-configuration of Figure 1, a square or rectangular configuration can also be utilized.

In the second step of the method according to the present invention, the DCT coefficients of each bit plane are scanned according to the sub-groups provided by the first step. It should be noted that the term "scanning" refers to the order in which the DCT
10 coefficients are encoded. As can be seen from Figure 2, the scanning according to the present invention is performed at four levels.

The first level of scanning occurs at the bit-plane level 74. This means that one bit-plane of the all of the DCT coefficients are encoded before the next bit-plane. The second level of scanning occurs at the sub-group level 76. This means that one sub-group is
15 completely encoded before the next sub-group is encoded in each of the bit planes.

The third level of scanning occurs at the block level 78. This refers to the order in which the DCT coefficient blocks are encoded. As previously described, the DCT coefficient blocks included in each bit-plane will be arranged into similar sub-groups. Thus, scanning at the block level 78 refers to the order in which the same sub-group of DCT
20 coefficients from each of the blocks is encoded. In some situations, it may be preferable to encode the sub-groups from each block in spatial order. The fourth level of scanning is at the coefficient level 80. This refers to the order in which the individual coefficients from each block are encoded.

An example of the DCT coefficient scanning (second step) according to the present invention is shown in Figure 3. In order to simplify illustration, only four DCT
25 coefficient blocks of a bit-plane is shown and only scanning of the first three sub-groups is shown. Further, the numbers included in the DCT coefficient blocks indicate the order in which the coefficients are encoded. As can be seen, all of the scanning occurs within the same bit-plane, which implies scanning at the bit-plane level 74. Further, one sub-group is
30 completely scanned before the next sub-group, which implies scanning at the sub-group level 76.

As can be further seen from Figure 3, the scanning at the block level 78 occurs in spatial order. However, the present invention contemplates other orders. For example, if an image is at the center of a picture, it would be preferable to start scanning with a DCT

coefficient block corresponding to the center of the picture. As is further evident, scanning at the coefficient level 80 in the second sub-group of coefficients occurs from top to bottom and left to right. Also, scanning at the coefficient level 80 in the third sub-group of coefficients occurs in a diagonal pattern, as shown. This demonstrates that the present invention contemplates scanning at the coefficient level 80 in a number of predetermined patterns.

The foregoing description of the present invention have been presented for the purposes of illustration and description. It is not intended to limit the invention to the precise forms disclosed. Many modifications and variations are possible in light of the above teachings. Therefore, it is not intended that the scope of the invention should be limited by the detail description.

CLAIMS:

1. A method for scanning discrete cosine transform (DCT) coefficients, the method comprising the steps of:
arranging the DCT coefficients into sub-groups; and
scanning the DCT coefficients according to the sub-groups.
5
2. The method of claim 1, wherein each of the sub-groups has an L-configuration.
3. The method of claim 1, wherein each of the sub-groups is scanned in a spatial
10 order.
4. The method of claim 1, wherein the scanning of the DCT coefficients occurs bit-plane by bit-plane.
- 15 5. The method of claim 1, wherein the DCT coefficients are enhancement layer DCT coefficients.
6. A method for encoding video data, the method comprising the steps of:
transforming the video data into a plurality of DCT coefficients;
20 arranging the DCT coefficients into sub-groups; and
scanning the DCT coefficients according to the sub-groups.
7. The method of claim 6, wherein each of the sub-groups has an L-configuration.
25
8. The method of claim 6, wherein each of the sub-groups is scanned in a spatial order.

9. The method of claim 6, wherein the scanning of the DCT coefficients occurs bit-plane by bit-plane.
10. The method of claim 6, wherein the video data is enhancement layer video
5 data.

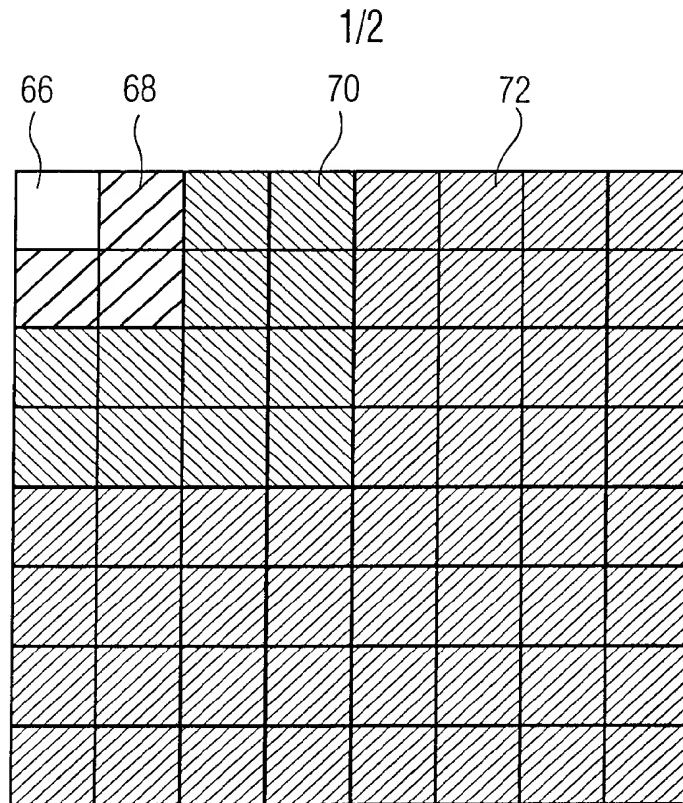


FIG. 1

```

FOR EACH BIT-PLANE BP 74
DO
  FOR EACH SUB-GROUP SG 76
  DO
    FOR EACH BLOCK B 78
    DO
      FOR EACH COEFFICIENT IN SUB-GROUP SG 80
      DO
        PROCESS BIT-PLANE BP OF THIS
        COEFFICIENT IN SUB-GROUP SB
        OF BLOCK B
      DONE
    DONE
  DONE
DONE
  
```

FIG. 2

1	5	17	22					2	8	29	34				
6	7	21	23					9	10	33	35				
18	20	24	27					30	32	36	39				
19	25	26	28					31	37	38	40				
3	11	41	46					4	14	53	58				
12	13	45	47					15	16	57	59				
42	44	48	51					54	56	60	63				
43	49	50	52					55	61	62	64				

FIG. 3

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 00/12247

A. CLASSIFICATION OF SUBJECT MATTER IPC 7 H04N7/24 H04N7/26 H04N7/46 H04N7/50 H04N7/36 H04N7/30		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 H04N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, PAJ, INSPEC		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 53613 A (MOTOROLA INC) 26 November 1998 (1998-11-26) the whole document ---	1-10
X	US 5 528 299 A (DUFOR CECILE ET AL) 18 June 1996 (1996-06-18) column 7, line 55 -column 8, line 25 figure 2B ---	1-3,6-8
A	US 5 109 451 A (AONO TOMOKO ET AL) 28 April 1992 (1992-04-28) column 2, line 52 -column 3, line 19 figures 3A-10B --- <div style="text-align: right;">-/--</div>	1-4,6-9
<div style="display: flex; justify-content: space-between;"> <input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex. </div>		
* Special categories of cited documents :		
<div style="display: flex; justify-content: space-between;"> <div style="width: 48%;"> <p>*A* document defining the general state of the art which is not considered to be of particular relevance</p> <p>*E* earlier document but published on or after the international filing date</p> <p>*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>*O* document referring to an oral disclosure, use, exhibition or other means</p> <p>*P* document published prior to the international filing date but later than the priority date claimed</p> </div> <div style="width: 48%;"> <p>*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.</p> <p>*Z* document member of the same patent family</p> </div> </div>		
Date of the actual completion of the international search <div style="text-align: center; font-weight: bold;">15 March 2001</div>		Date of mailing of the international search report <div style="text-align: center; font-weight: bold;">22/03/2001</div>
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer <div style="text-align: center; font-weight: bold;">Fassnacht, C</div>

INTERNATIONAL SEARCH REPORT

Int ional Application No
PCT/EP 00/12247

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	LAURANCE N K ET AL: "EMBEDDED DCT CODING WITH SIGNIFICANCE MASKING" MULTIDIMENSIONAL SIGNAL PROCESSING, NEURAL NETWORKS. MUNICH, APR. 21 - 24, 1997, LOS ALAMITOS, CA: IEEE COMP. SOC. PRESS, US, 1997, pages 2717-2720, XP000867104 ISBN: 0-8186-7920-4 section 2.1. ----	1,4,6,9
P,X	VAN DER SCHAAR M ET AL: "Embedded DCT and wavelet methods for fine granular scalable video: analysis and comparison" IMAGE AND VIDEO COMMUNICATIONS AND PROCESSING 2000, SAN JOSE, CA, USA, 25-28 JAN. 2000, vol. 3974, pages 643-653, XP000981435 Proceedings of the SPIE - The International Society for Optical Engineering, 2000, SPIE-Int. Soc. Opt. Eng, USA ISSN: 0277-786X abstract section 5. -----	1-10

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 00/12247

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9853613 A	26-11-1998	NONE	
US 5528299 A	18-06-1996	FR 2667753 A	10-04-1992
		FR 2670646 A	19-06-1992
		FR 2678126 A	24-12-1992
		FR 2679092 A	15-01-1993
		CA 2070827 A	10-04-1992
		DE 69116128 D	15-02-1996
		DE 69116128 T	01-08-1996
		EP 0504376 A	23-09-1992
		FI 922648 A	08-06-1992
		JP 5503403 T	03-06-1993
		WO 9206563 A	16-04-1992
US 5109451 A	28-04-1992	JP 1276980 A	07-11-1989
		JP 1276981 A	07-11-1989
		JP 2017517 C	19-02-1996
		JP 7048858 B	24-05-1995
		EP 0339589 A	02-11-1989